University of Information Technology – Vietnam National University HCMC
Faculty of Computer Science

CS431.N21 – Deep Learning and Applications

Analysis of FCN for Semantic Segmentation

Trần Thái Bảo – 20520410
Lê Đoàn Phúc Minh – 20520243
Nguyễn Duy Đạt – 20520435
Trần Hữu Khoa – 20520222

Introduction

Introduction

Image Segmentation is a broad part of Computer Vision, in image segmentation we classify every pixel of the image into one of the classes.

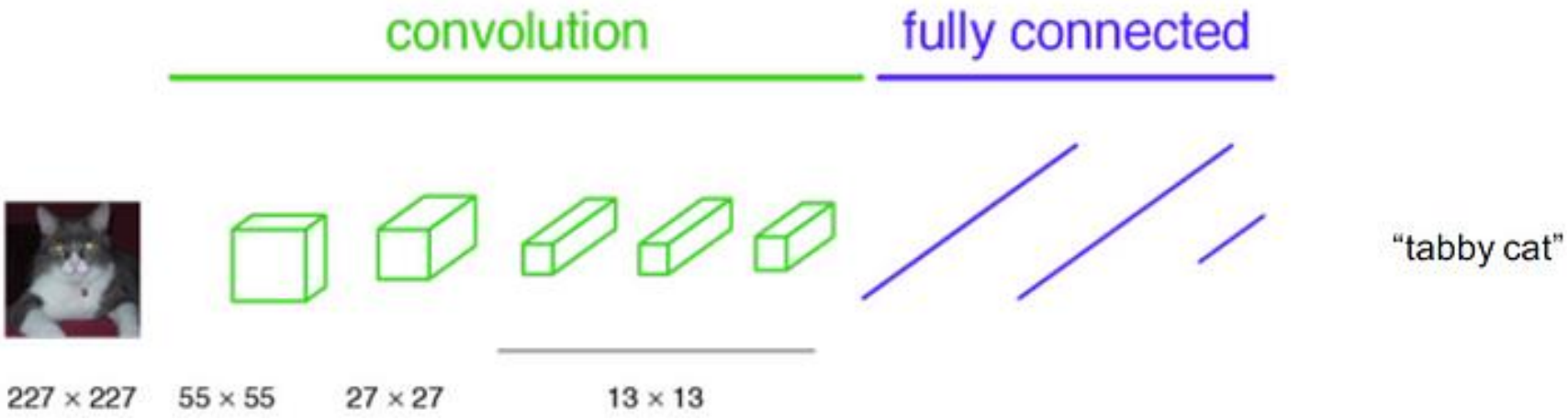
Type of Segmentation

Semantic segmentation

Instance segmentation

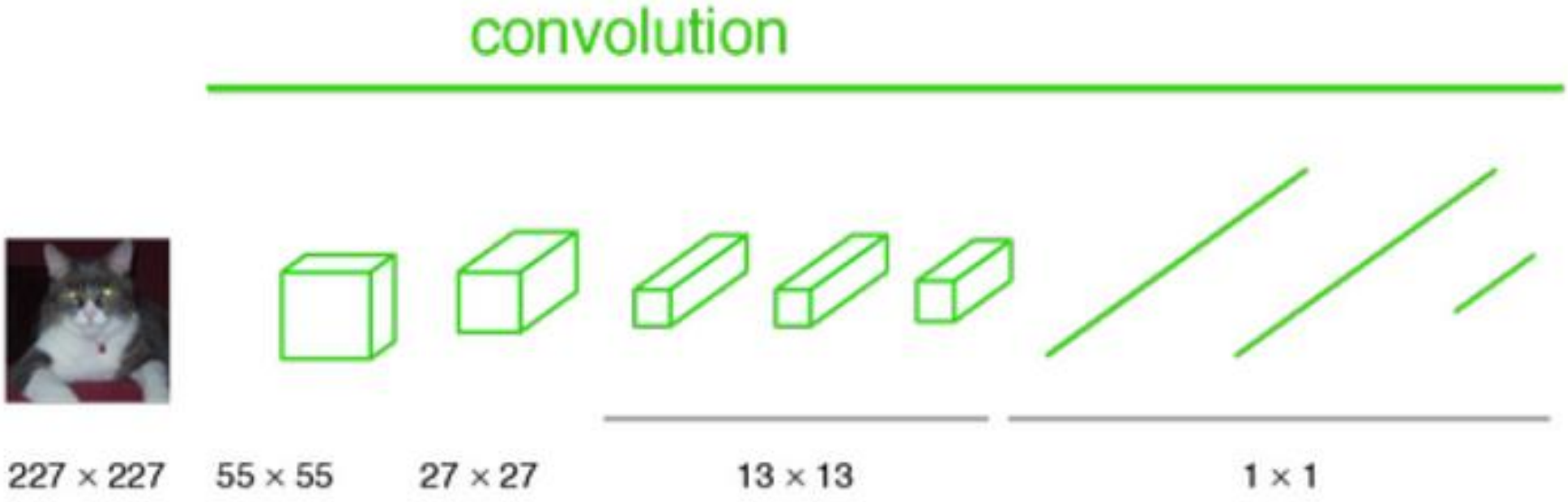
Fully Convolutional Networks (FCN)

Most of us may familar with CNNs in Image Classification



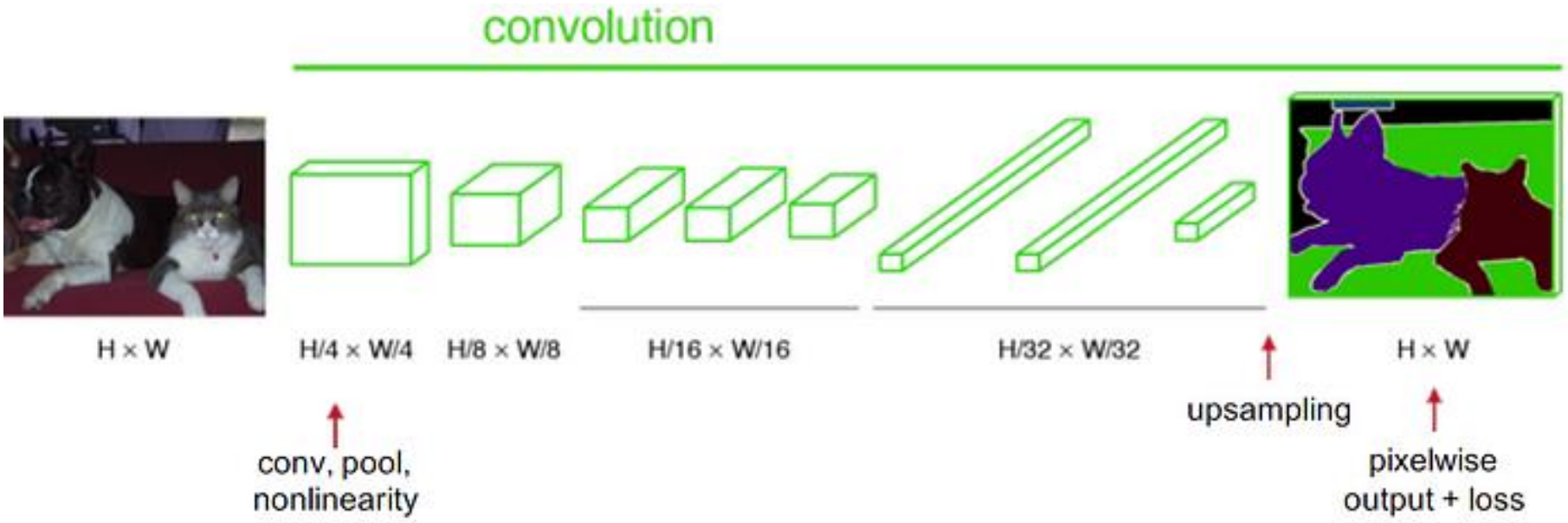
Architecture of a CNN model

What if...?



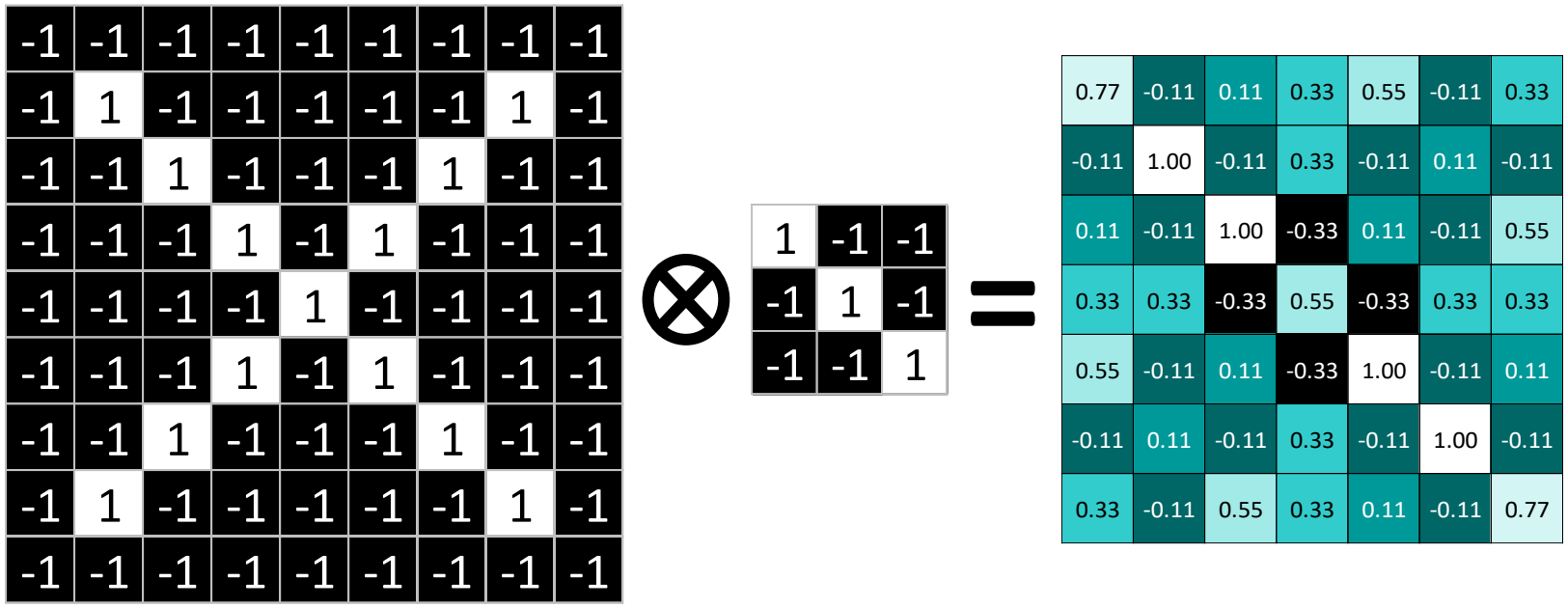
CNN - From Image Classification to Semantic Segmentation

CNN - From Image Classification to Semantic Segmentation

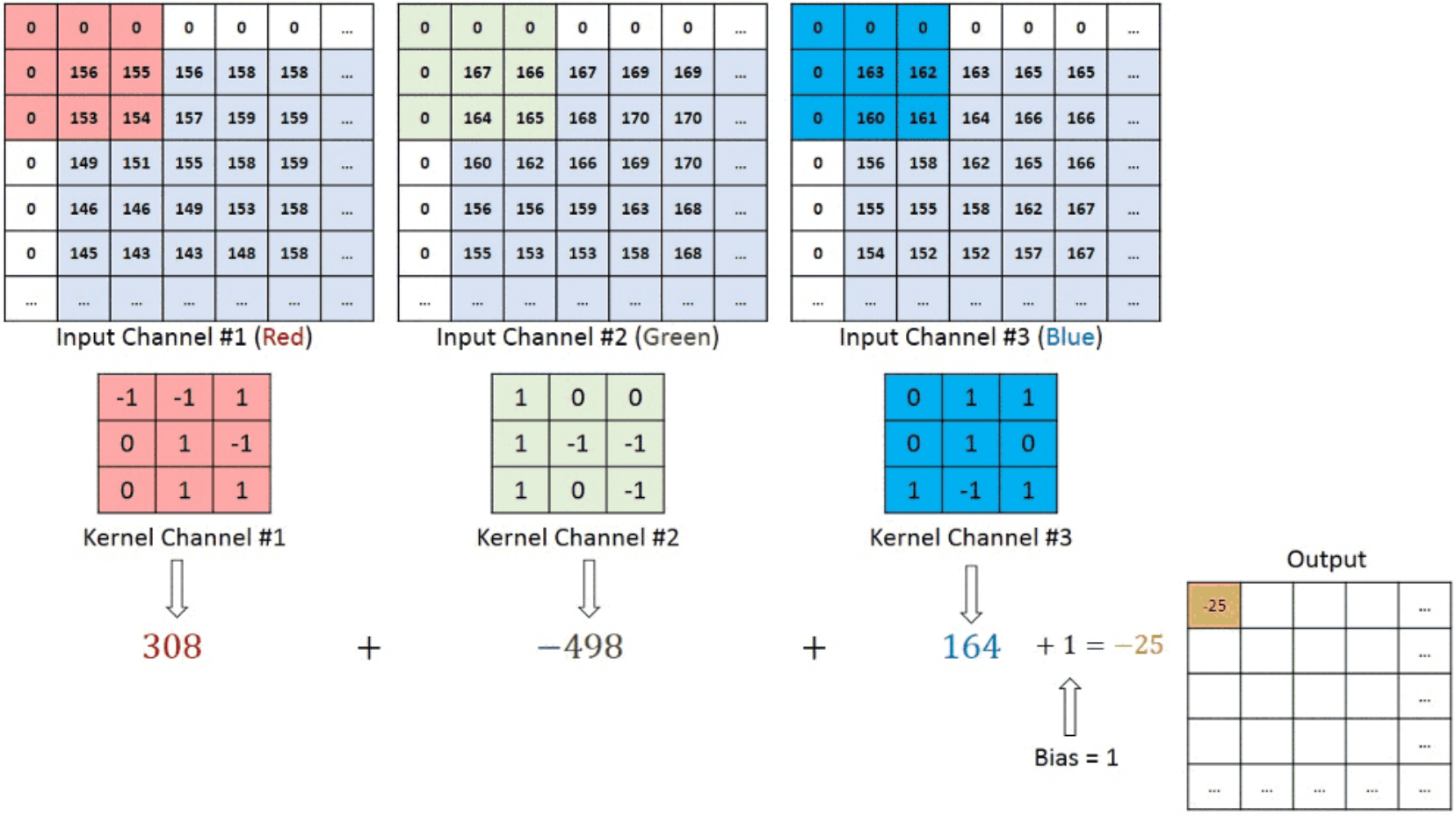


Convolution Matrix

Convolution Matrices are used to extract spatial features

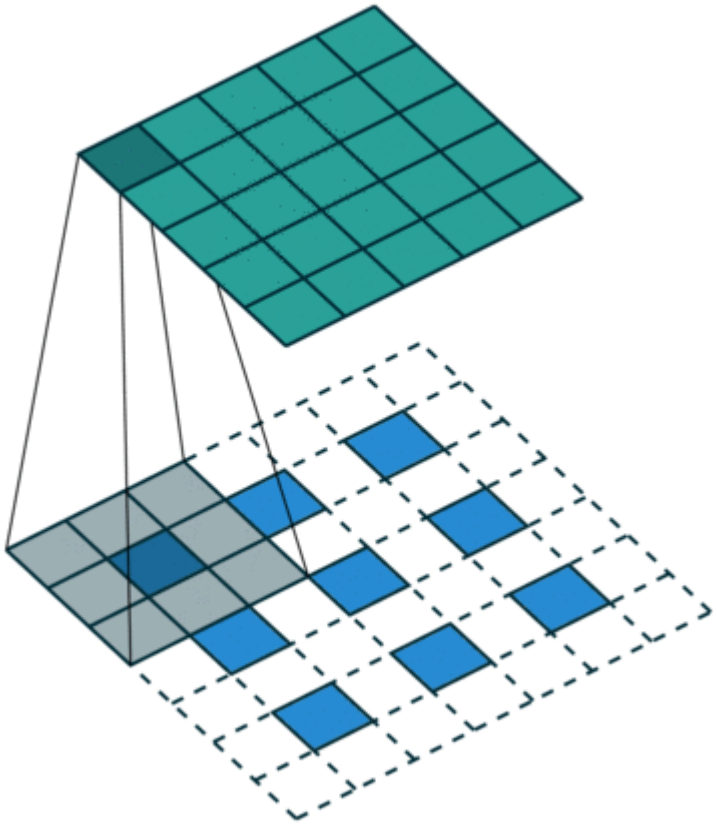


Convolution Layer



DeConvolution

DeConvolution is a technique to upsample the image by using Convolution Matrices.



Why DeConvolution ?

10	20
30	40

→

10	10	20	20
10	10	20	20
30	30	40	40
30	30	40	40

Nearest Neighbor

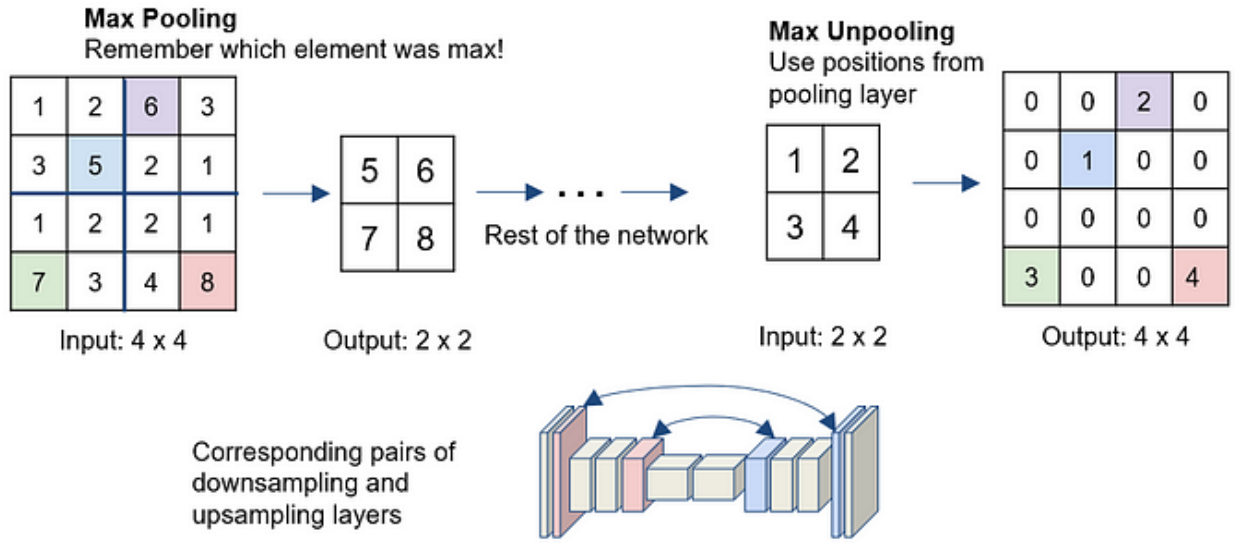
10	20
30	40

→

10	12	17	20
15	17	22	25
25	27	32	35
30	32	37	40

Bi-Linear Interpolation

Unpooling



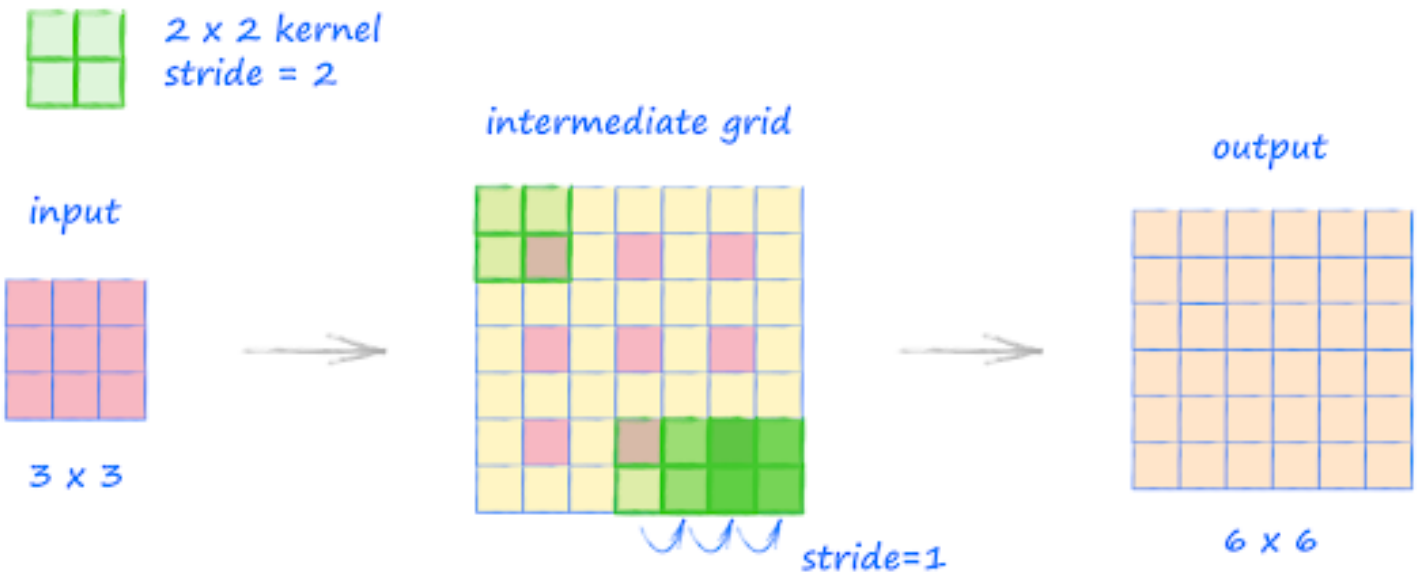
Why DeConvolution ?

Problems with these upscaling techniques ?

They do not learn from data

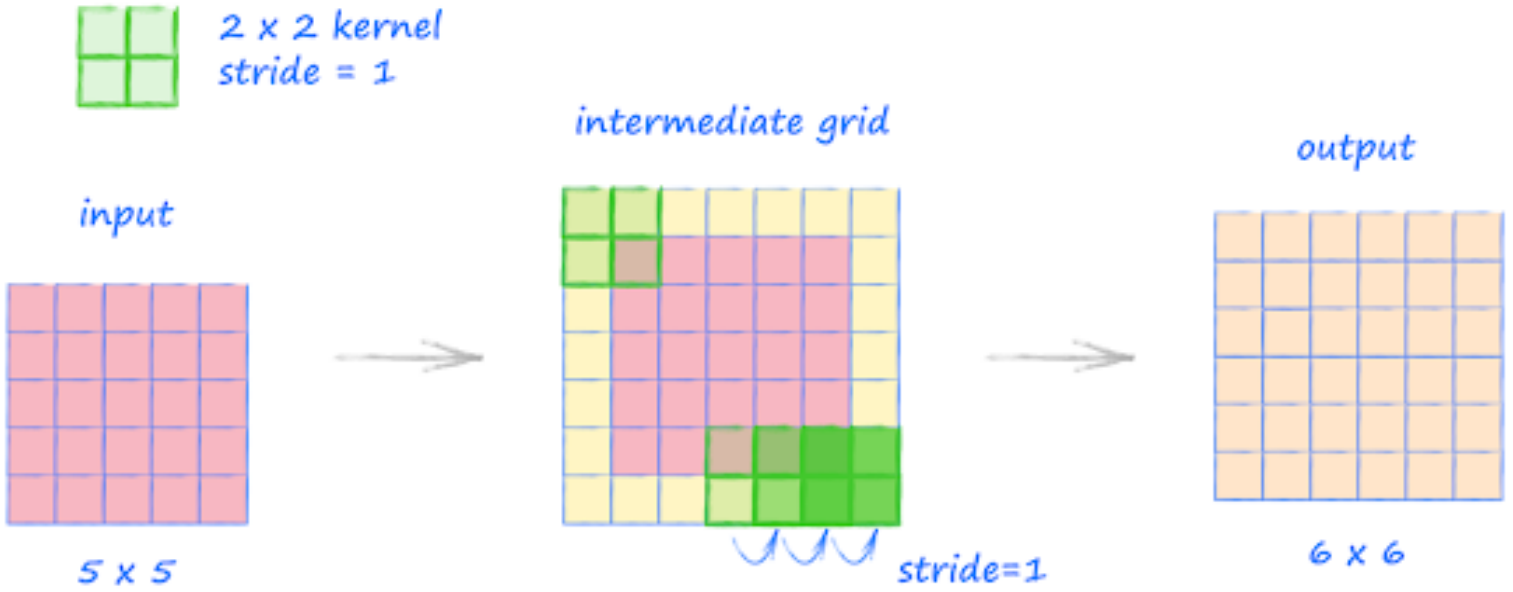
DeConvolution - Procedure

- **Step 1:** Apply the transformation to input (Add 0-cells between rows and cols)
- **Step 2:** Do convolution to input with given padding and stride informations



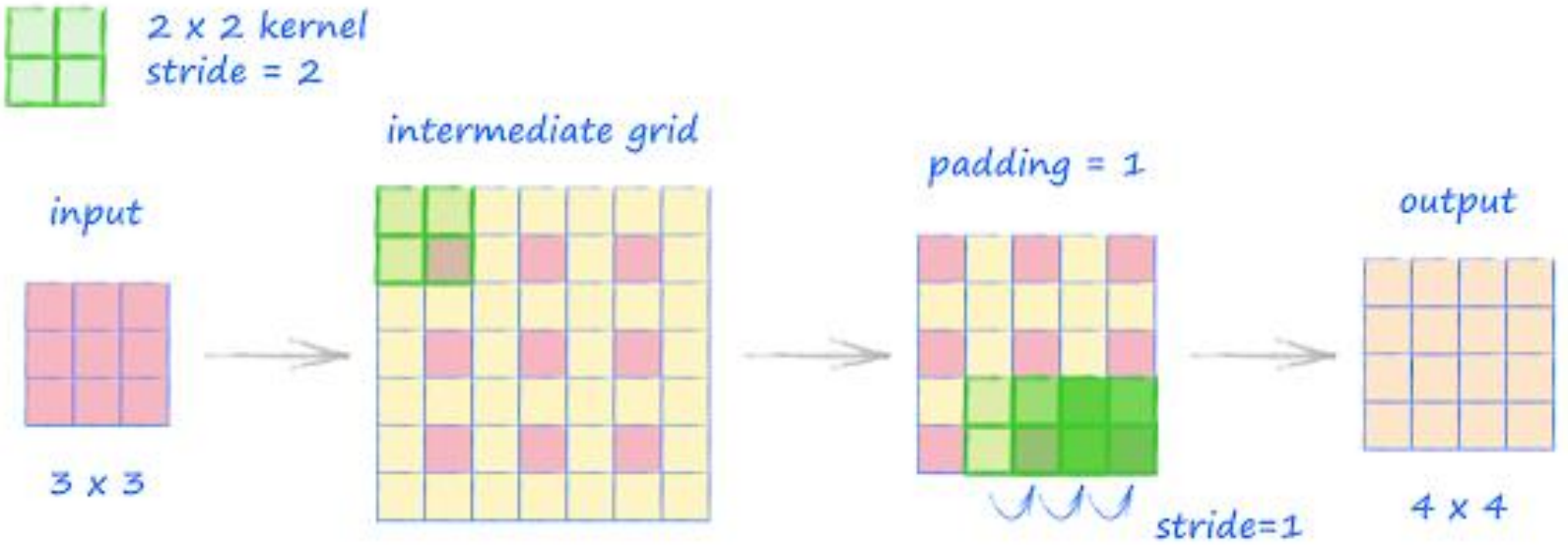
Stride 2, No Padding

DeConvolution – Other examples



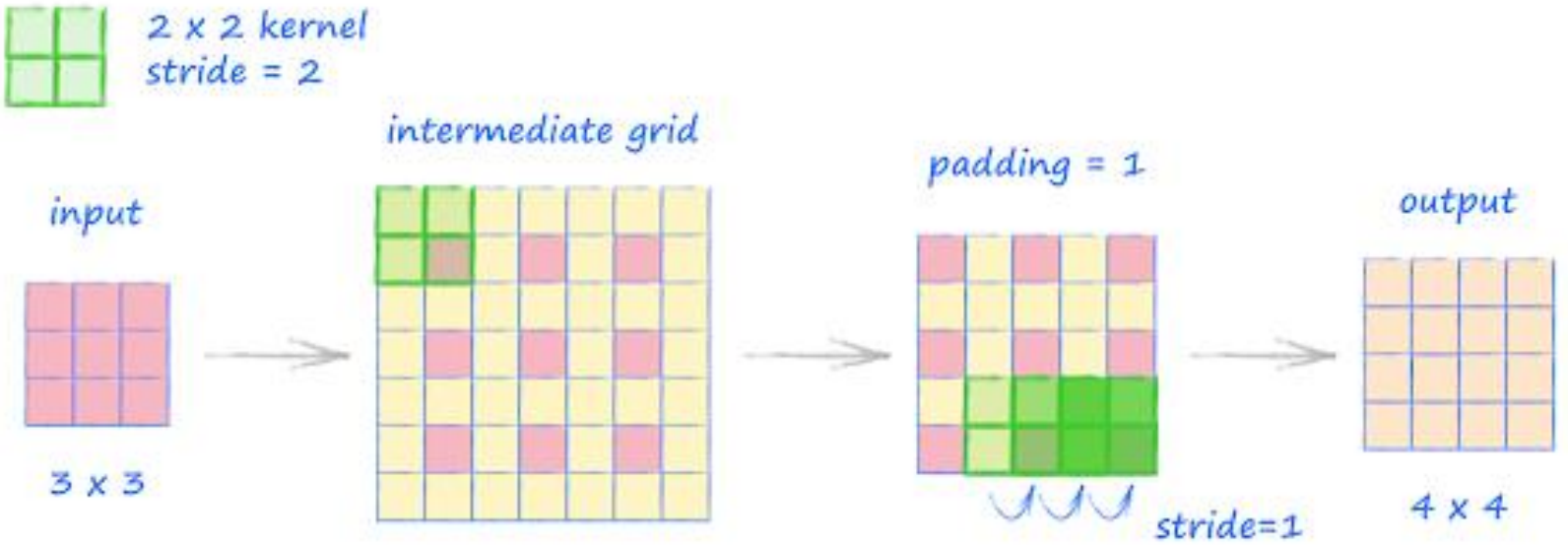
Stride 1, No Padding

DeConvolution – Other examples



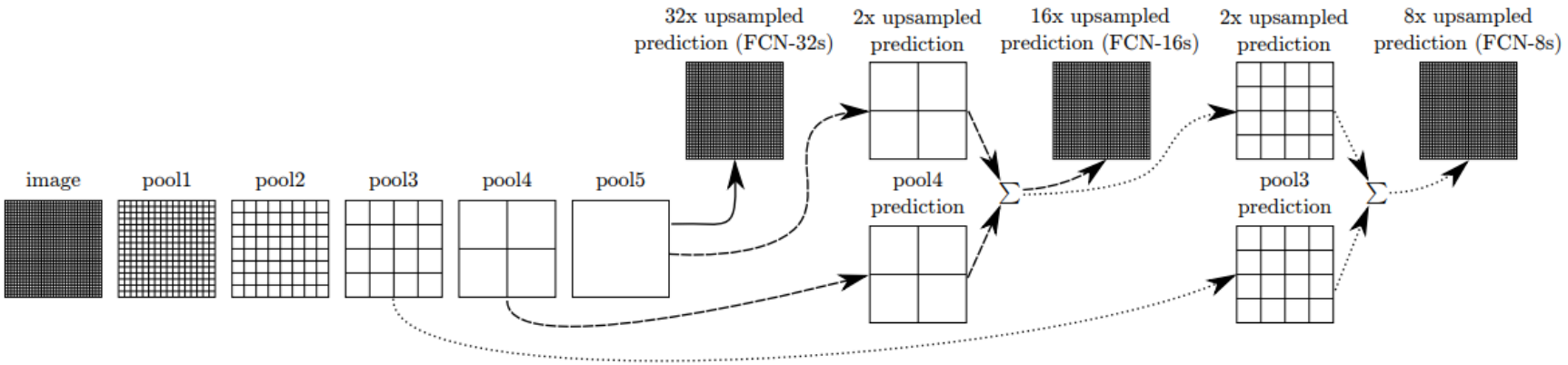
Stride 2, Padding

DeConvolution – Other examples

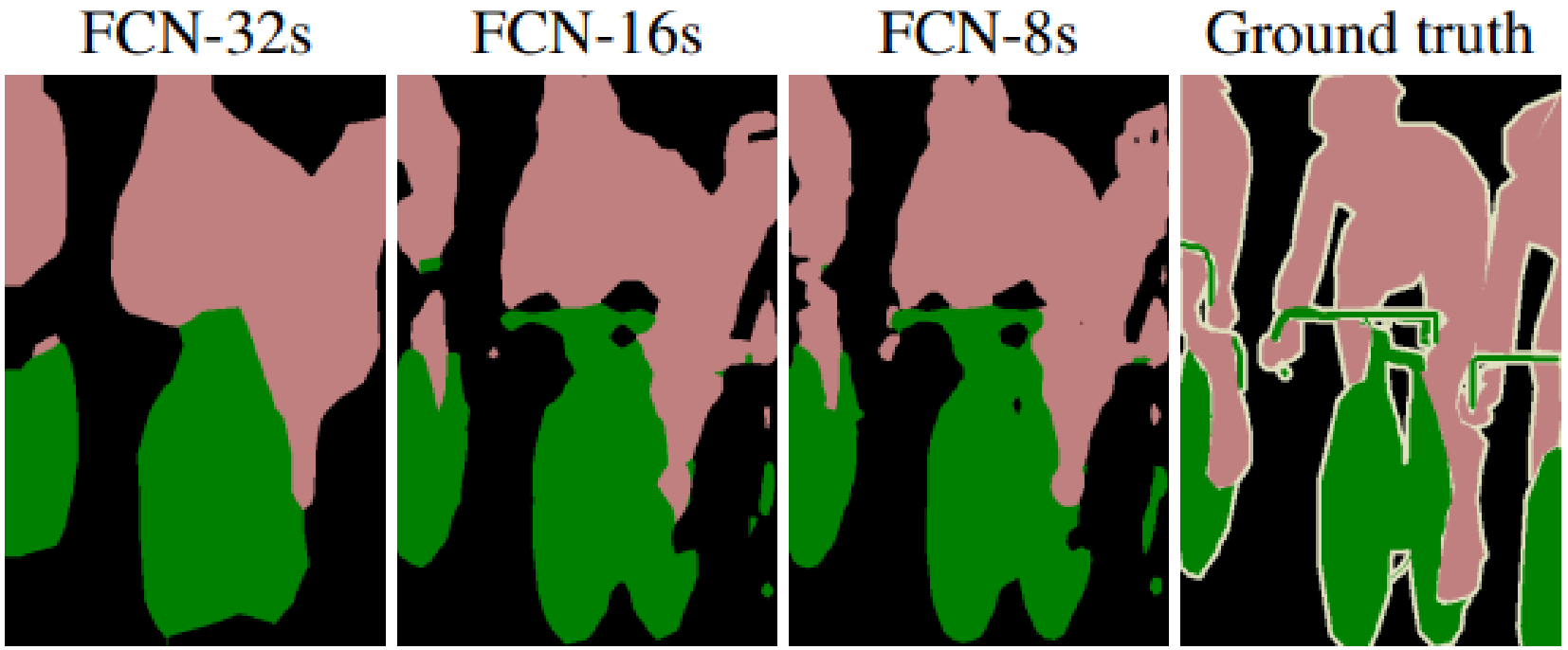


Stride 2, Padding

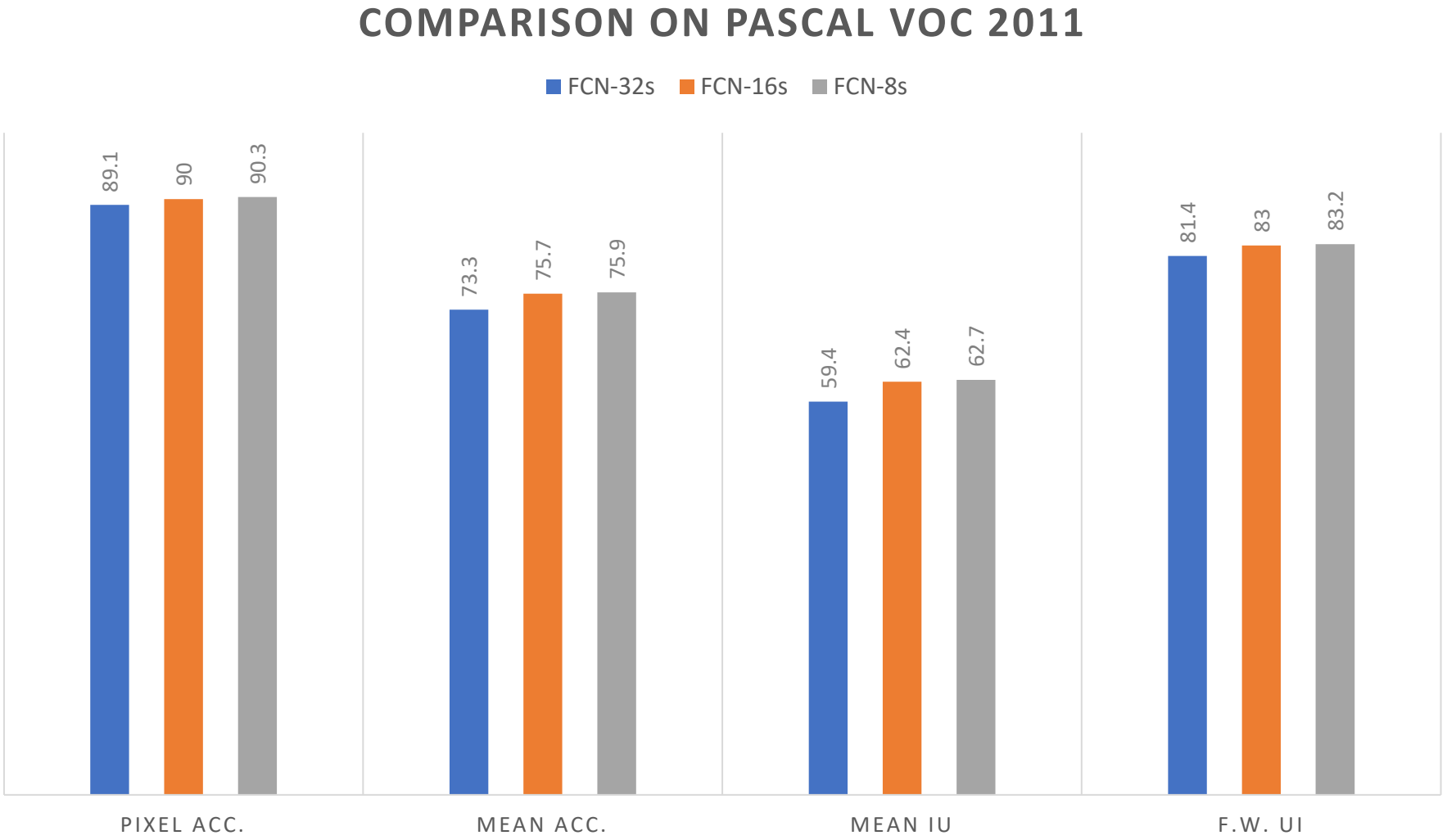
Architecture of FCN (32s/16s/8s)



Comparison between FCNs



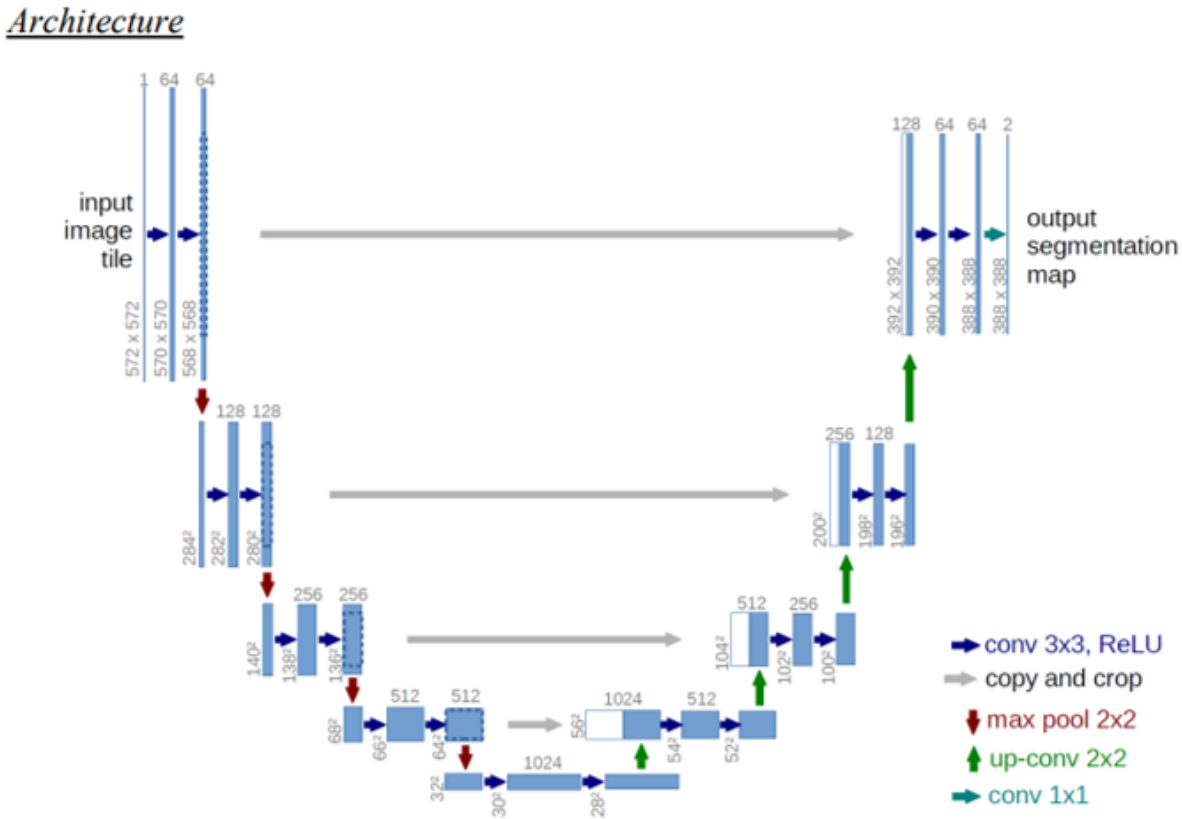
Comparison between FCNs



U-Net

Unet

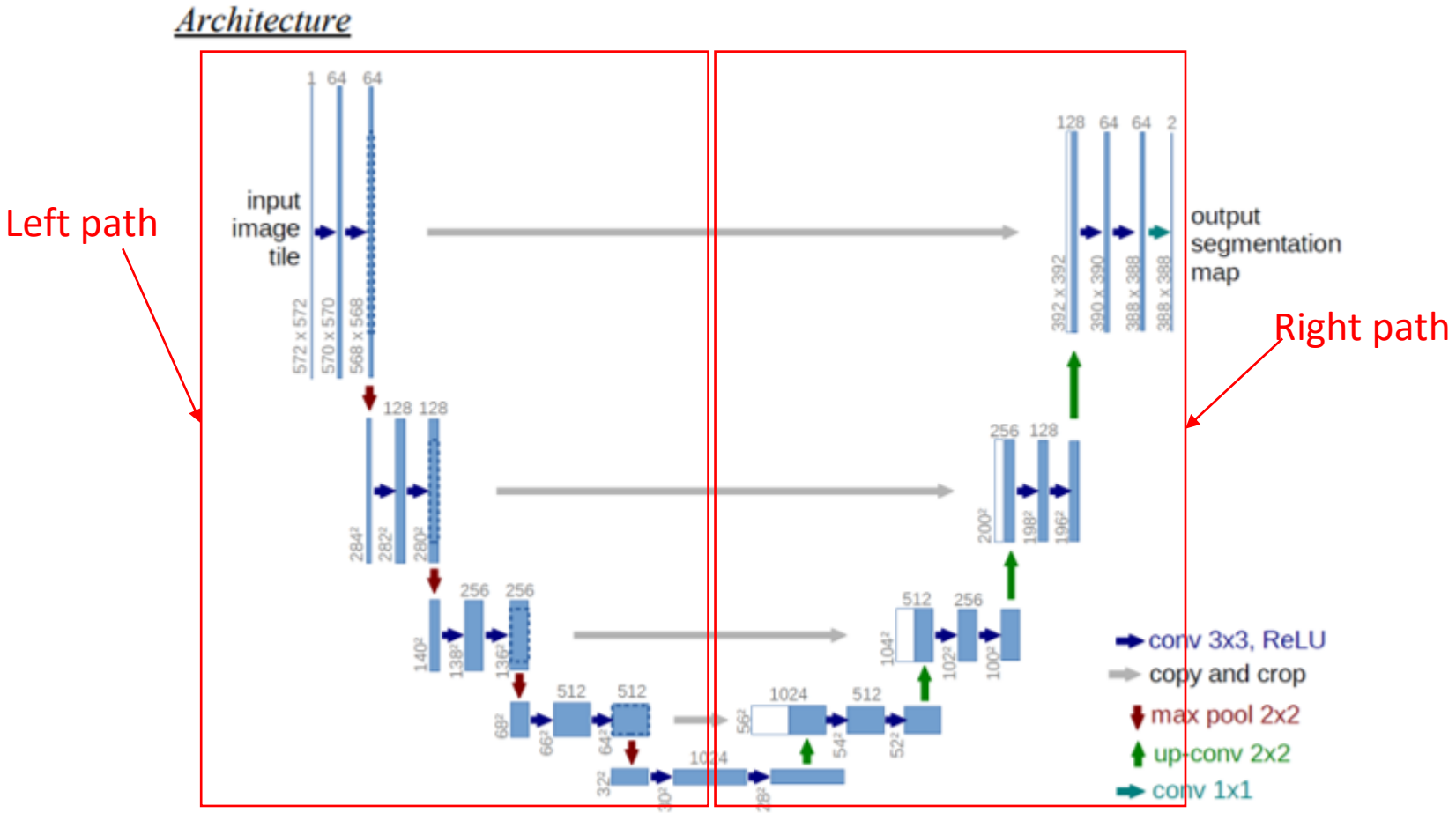
U-Net is a convolutional neural network originally developed for segmenting biomedical images. When visualized its architecture looks like the letter U and hence the name U-Net.



Unet

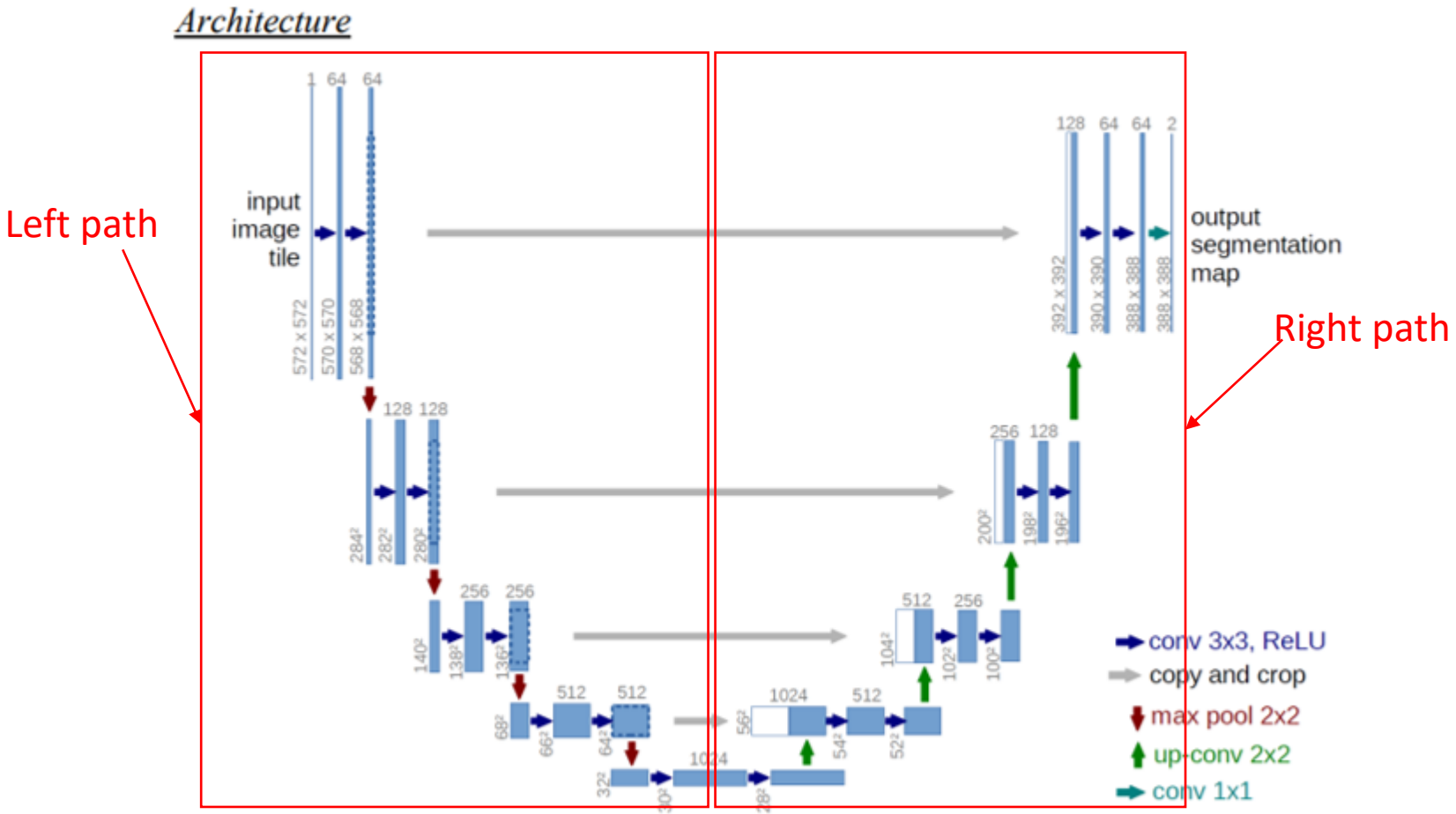
Unet's architecture is made up of two parts:

- The left path: contracting path
- The right path: expansive path



Unet

- Contracting path: capture context
- Expansive path: aid in precise localization

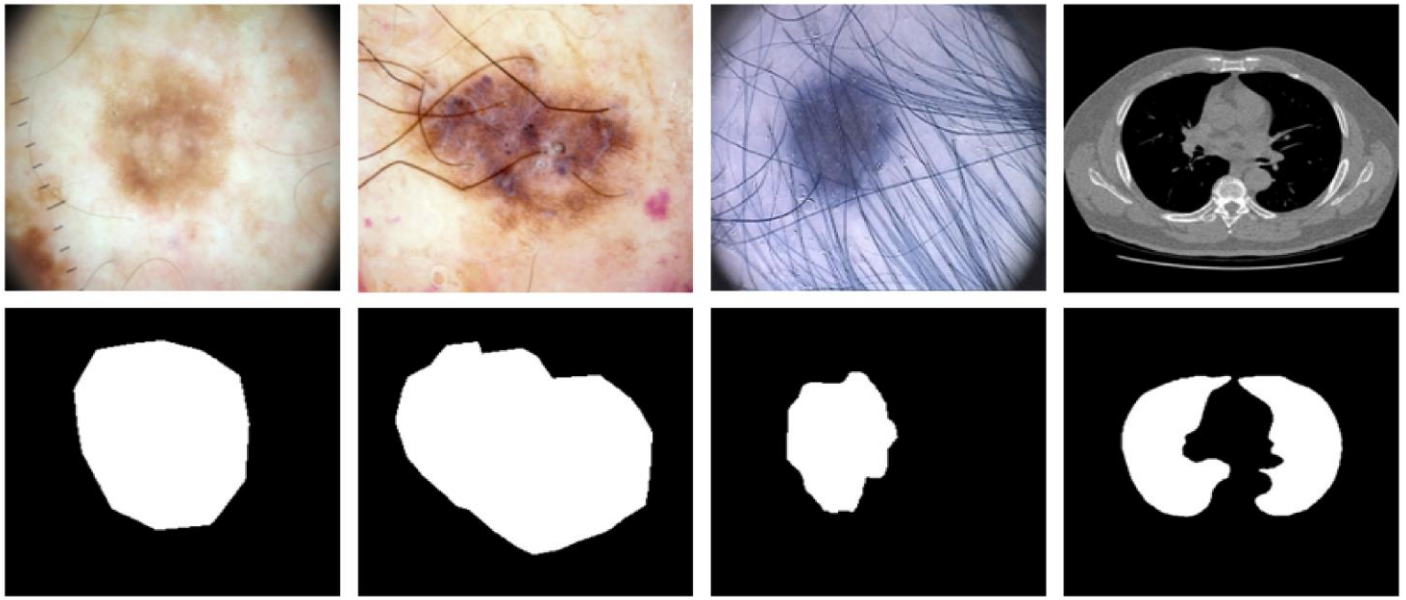


Unet

- With Unet in segmentation, we need to reconstruct the image from the feature vector created by CNN.
- We convert feature map into a vector and reconstruct an image from this vector.
- Unet has no any fully connected layers. The features are connected by the right path, so it don’t need fully connected network.

Application

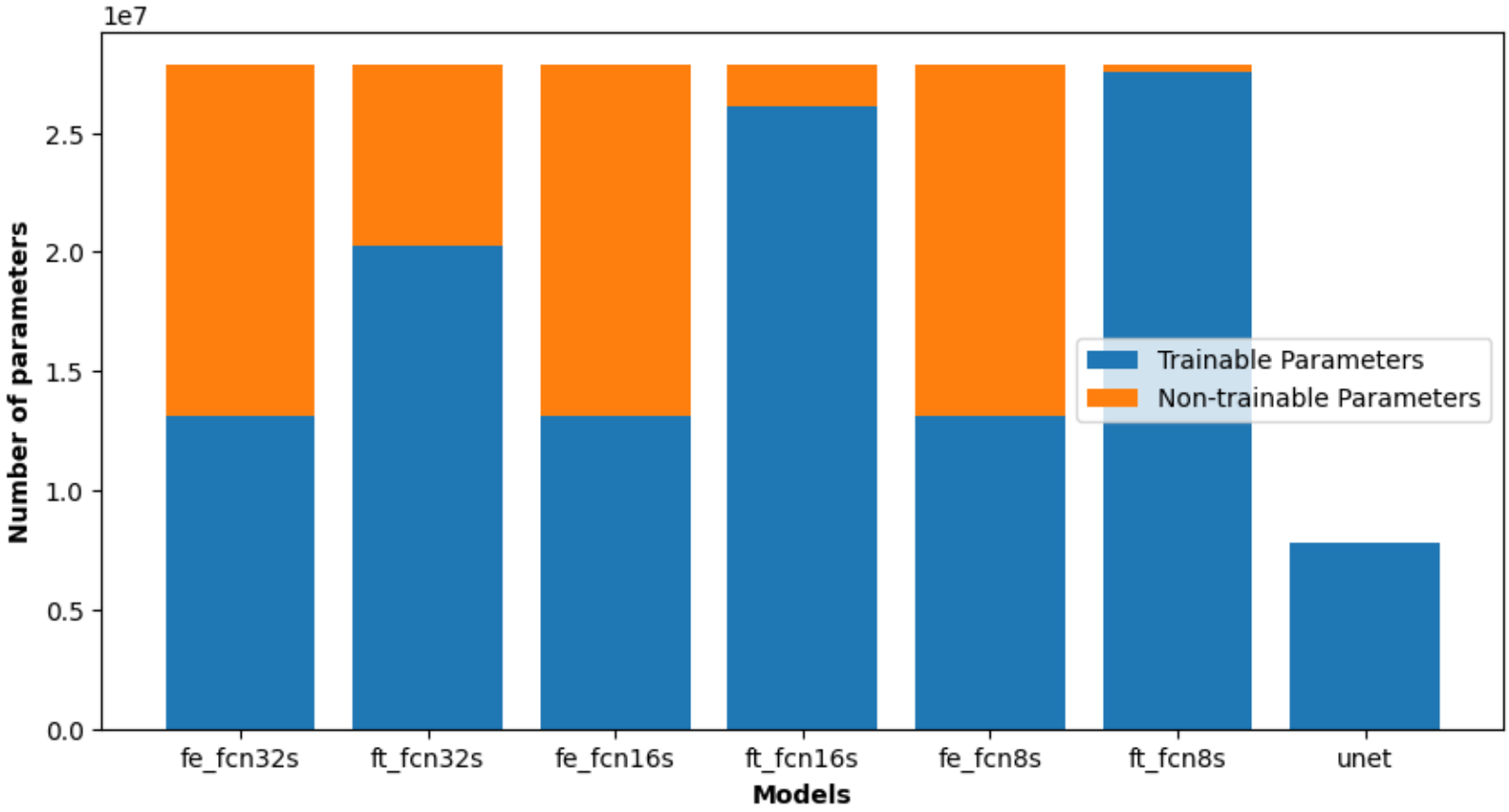
- Medical imaging
- Object detection
- Traffic control systems
- Machine vision
- Content-based image retrieval



Experiments

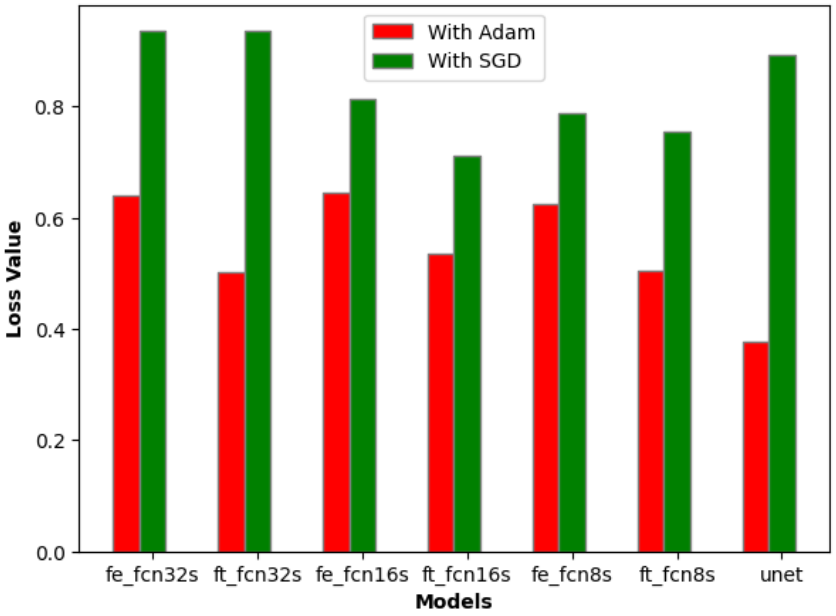
Total parameters

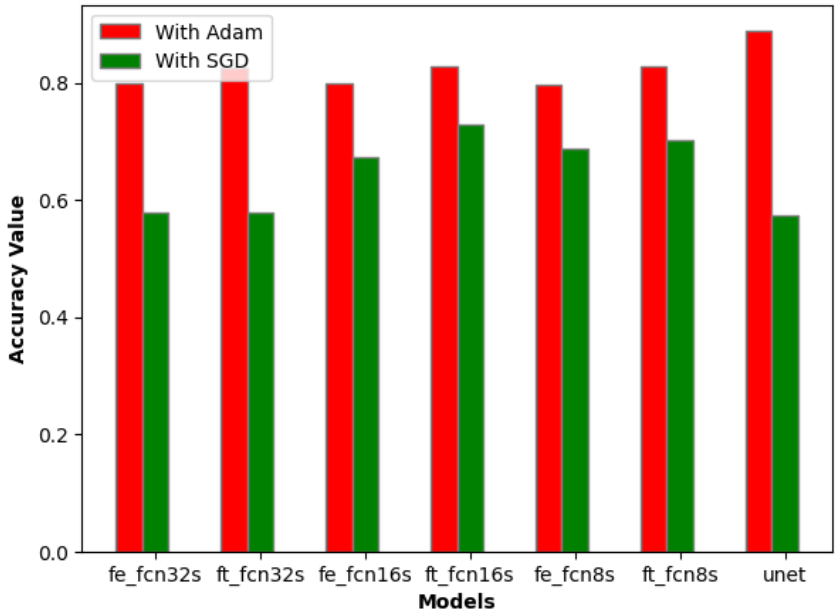
- The comparison about total parameters among the types of models



Loss and Accuracy

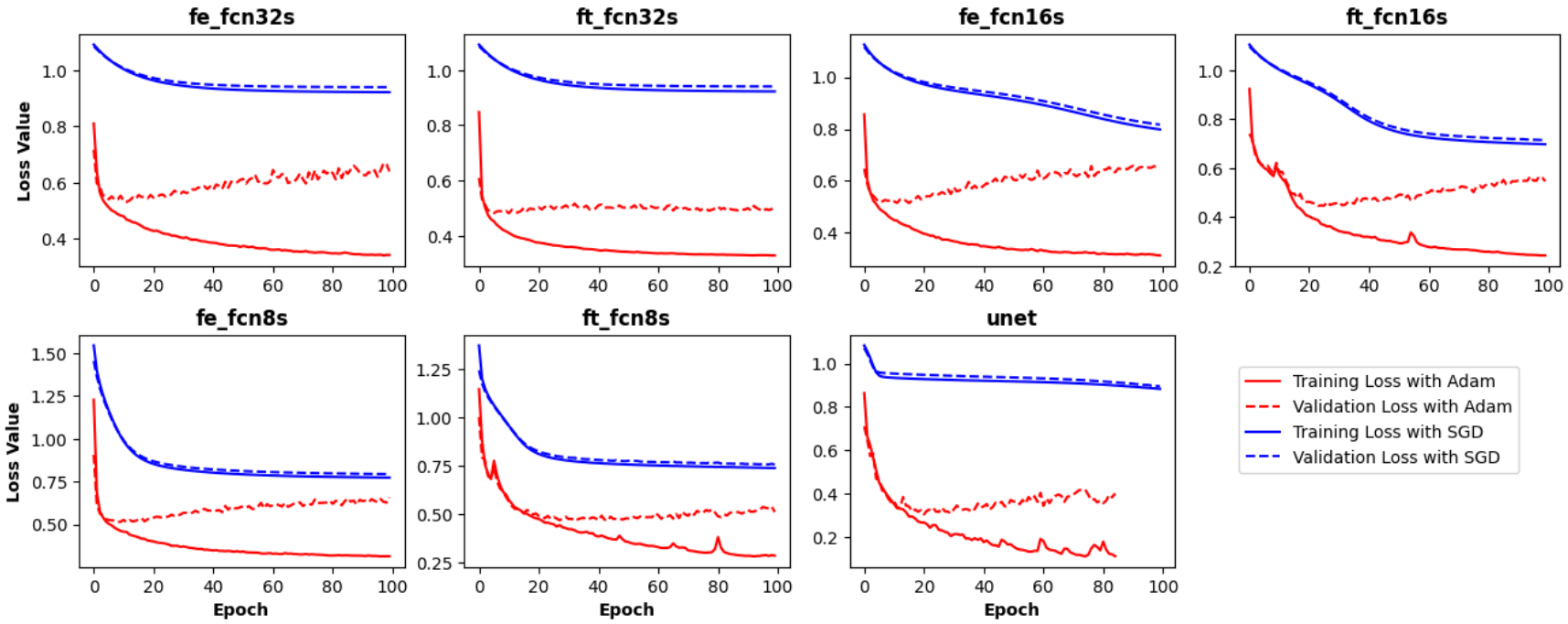
- The comparison of loss and accuracy





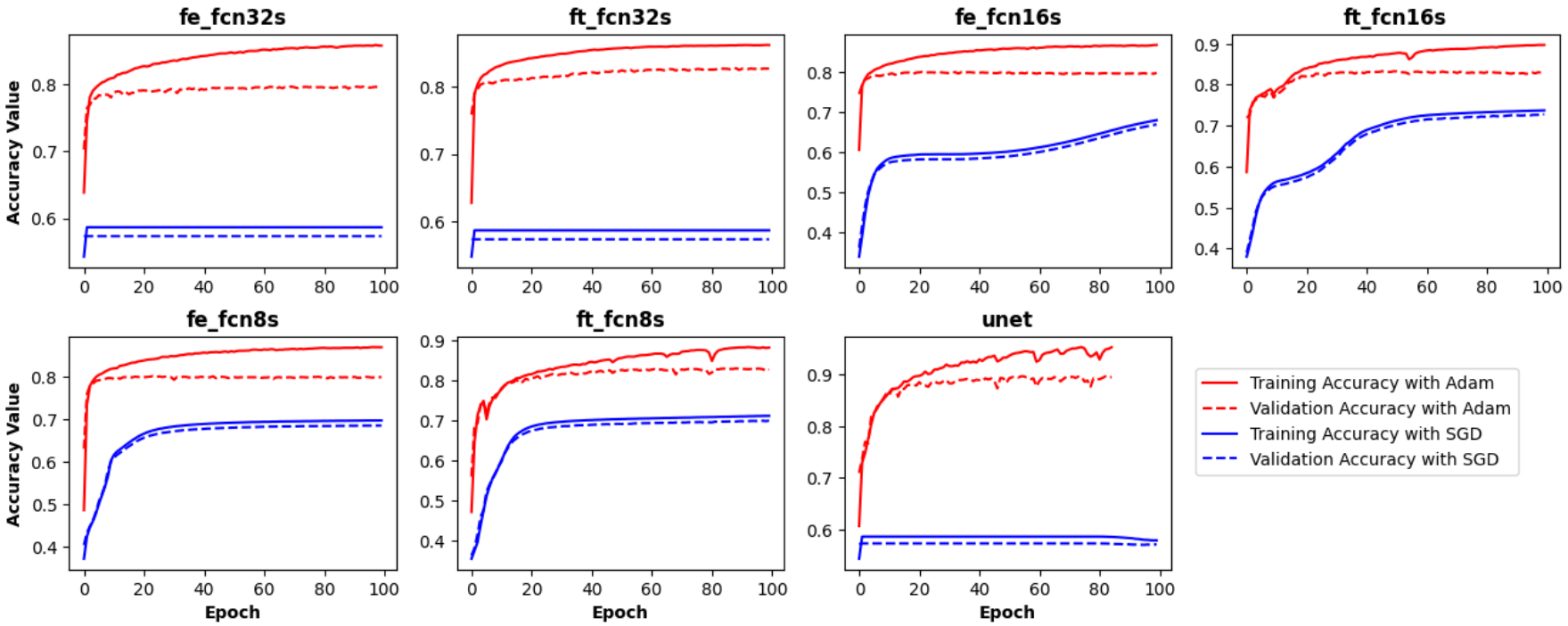
Loss convergence speed

- The comparison of loss convergence speed



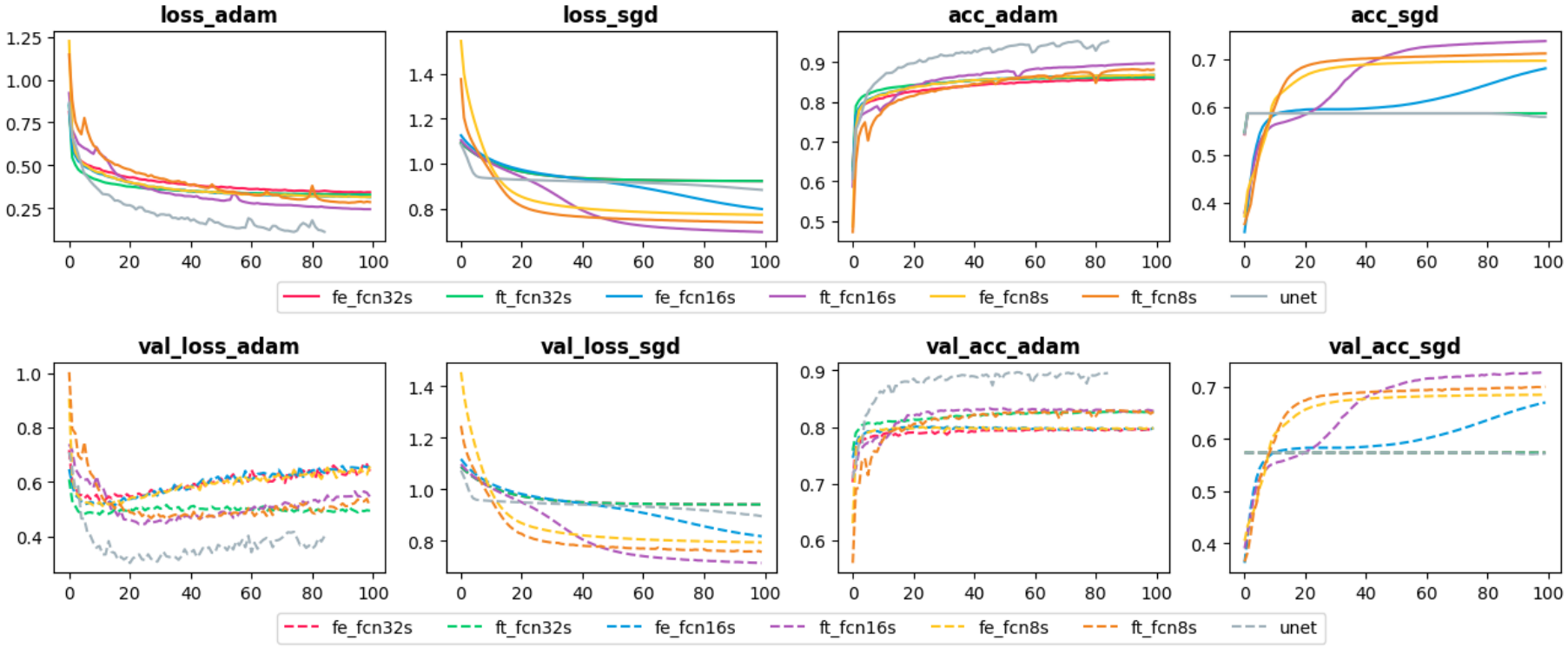
Accuracy convergence speed

- The comparison of accuracy convergence speed



All convergences

- The comparison of all convergences



Comparison the predictions

- The comparison of predictions

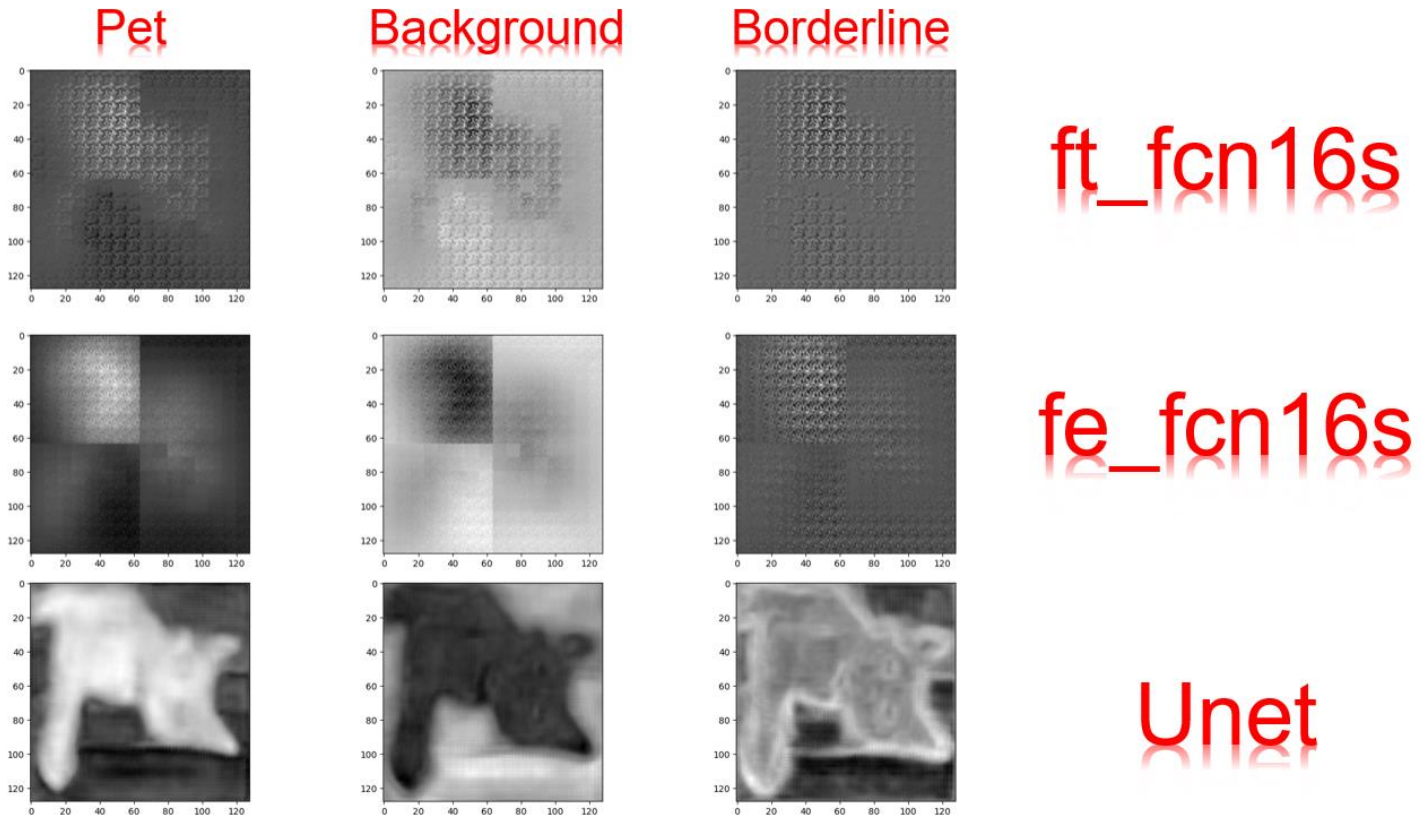
ft_fcn16s

fe_fcn16s

Unet

Comparison the feature maps

- The comparison of feature maps



Reference

Unet & FCN Demo:

https://colab.research.google.com/drive/1iisL6plYsJCXE21KPNvDZD5uQucmX1tt?usp=sharing

https://colab.research.google.com/drive/1e_55b_8aiam9h_P-NPuNkQXZL_LZGsZq

https://colab.research.google.com/drive/1DGWbaS2lQcp1tAXizUQHVPsrMmkCvL2U?usp=sharing

Squeeze U-Net: A Memory and Energy Efficient Image Segmentation Network (thecvf.com)

[1505.04597] U-Net: Convolutional Networks for Biomedical Image Segmentation (arxiv.org)